1. 准备数据

三个数据文件：

```
users.dat
movies.dat
ratings.dat
```

2. 上传到 HDFS

使用 HDFS 用户来执行：

```
hadoop fs -mkdir -p /moviedata/
hadoop fs -put ~/users.dat /moviedata/
hadoop fs -put ~/movies.dat /moviedata/
hadoop fs -put ~/ratings.dat /moviedata/
```

3. 建表

使用 hive 用户，进入到 hive 交互式操作界面执行建表导入数据等操作：

```
drop table if exists users;
create table users(UserID BigInt, Gender String, Age Int, Occupation String, Zipcode String) row format serde
'org.apache.hadoop.hive.serde2.RegexSerDe' with serdeproperties('input.regex'='(.*)::(.*)::(.*)::(.*)::
(.*)','output.format.string'='%1$s %2$s %3$s %4$s %5$s') stored as textfile;

drop table if exists movies;
create table movies(MovieID BigInt, Title String, Genres String) row format serde
'org.apache.hadoop.hive.serde2.RegexSerDe' with serdeproperties('input.regex'='(.*)::(.*)::
(.*)','output.format.string'='%1$s %2$s %3$s') stored as textfile;

drop table if exists ratings;
create table ratings(UserID BigInt, MovieID BigInt, Rating Double, Timestamped String) row format serde
'org.apache.hadoop.hive.serde2.RegexSerDe' with serdeproperties('input.regex'='(.*)::(.*)::(.*)::
(.*)','output.format.string'='%1$s %2$s %3$s %4$s') stored as textfile;
```

4. 导入数据

导入数据：

```
load data inpath "/moviedata/users.dat" into table users;
load data inpath "/moviedata/movies.dat" into table movies;
load data inpath "/moviedata/ratings.dat" into table ratings;
```

5. 查询数据

查询导入是否成功：

```
select * from users limit 3;
select * from movies limit 3;
select * from ratings limit 3;
```